



Early Journal Content on JSTOR, Free to Anyone in the World

This article is one of nearly 500,000 scholarly works digitized and made freely available to everyone in the world by JSTOR.

Known as the Early Journal Content, this set of works include research articles, news, letters, and other writings published in more than 200 of the oldest leading academic journals. The works date from the mid-seventeenth to the early twentieth centuries.

We encourage people to read and share the Early Journal Content openly and to tell others that this resource exists. People may post this content online or redistribute in any way for non-commercial purposes.

Read more about Early Journal Content at <http://about.jstor.org/participate-jstor/individuals/early-journal-content>.

JSTOR is a digital library of academic journals, books, and primary source objects. JSTOR helps people discover, use, and build upon a wide range of content through a powerful research and teaching platform, and preserves this content for future generations. JSTOR is part of ITHAKA, a not-for-profit organization that also includes Ithaka S+R and Portico. For more information about JSTOR, please contact support@jstor.org.

THE AMERICAN JOURNAL OF SOCIOLOGY

VOLUME XV

MAY, 1910

NUMBER 6

THE SOCIAL MARKING SYSTEM

PROFESSOR FRANKLIN H. GIDDINGS
Columbia University

In their quantitative aspect many data of sociology are positions in a scheme, rather than distances from zero. Tables of births, deaths, and migrations, like tables of height and of weight, give us measures from zero. Degrees of difference or of resemblance which we observe among our acquaintances and others, the affiliations of nationalities and of races, of religious beliefs, and of political interests, the values that we assign to ability and to conduct, and the social ranks that make up differentiated communities, are merely positions in a scheme.

Sir Francis Galton long ago showed that when positions in a scheme are successive, and may be successively numbered, and when the instances in which given phenomena fall into successively numbered positions, can be counted, and their frequencies set down, we can legitimately subject the numerical data so assembled to familiar methods of statistical analysis. We can plot their rough curves of magnitude and of frequency; we can ascertain their medians, quartiles, and probable errors; we can determine their modes. Whether we can also make significant use of their standard deviations and coefficients of variation, is a question that need not now be considered.

In order to assemble and to sort such data, however, we must have a marking-scale. Also we must have a concrete knowledge of the facts to be sorted, like that on which we rely when we assign numerical grades to examination papers, or to tests of

conduct. The marking-scale must not be arbitrary, a mere ingenious invention. It must be a simple and natural expression of observed relations. If possible, its elements should be derived from those common-sense gradings which the world uses in rough and ready fashion for the purposes of everyday life, and which therefore have the warrant of experience.

A familiar example of empirical gradings is the scale of consanguinity. The children of brothers or sisters we call first cousins, and describe them as one degree less nearly related than brothers or sisters. Children of first cousins we call second cousins, and describe them as two degrees less nearly related than brothers or sisters, and so on. These degrees, as everybody recognizes, are nothing more than successive positions in a scheme. No absolute numerical value, or distance from zero, attaches to any one of them. A kindred but less familiar example is offered in Lewis H. Morgan's ordinal arrangement of the successively wider and wider groupings composing tribal society. The arrangement is, namely: first grouping, the family; second grouping, the totem kin or clan; third grouping, the phratry or brotherhood of clans; fourth grouping, the tribe; fifth grouping, the federation of tribes.

One intent of this paper is to call attention to other rough marking-scales which we daily make use of in describing social phenomena, and to show how they may be extended and made sufficiently precise for scientific purposes. A second intent is to indicate how, by distributing a considerable mass of existing numerical data in accordance with the marking-scales to be described, we may, through the usual methods of statistical analysis, arrive at sociological conclusions that are interesting and possibly important.

To a great extent the phenomena of population are described by statistical measures from zero. Density, excess of births over deaths, emigration and immigration, are expressed in absolute numbers, of approximate accuracy. But such measures are not forthcoming when we ask a question like the following, namely, Is a population made up of native white Americans, foreign-born Irishmen and Germans, foreign-born Italians, Hungarians and

Poles, and both native and foreign negroes, more mixed, composite, or heterogeneous than one made up of native white Americans, Spanish-Americans, Chinamen, Hawaiian Islanders, and Indians not taxed, and, if so, how much more? To answer a question of this kind we must first assign each nationality or race to a position on a predetermined marking-scale of resemblance or affiliation. We must be able to say: Nationality E is so many points more unlike the native-born whites than Nationality B is, or to say: Race K is psychologically and culturally so many positions more distant from Race A than Race D is.

By general consent we have employed a scale of this kind for twenty years or more in our discussion of the immigration problem. No one ever questions the propriety of assigning the native-born of white parents to the initial position, which we will call "zero," nor of assigning the native-born of foreign-born whites to the second position which we will call "one." In like manner, without dispute the uncivilized races are assigned to the final position, whatever its number may happen to be. Assignments to intermediate positions are made by concurring judgments upon the varying degrees of resemblance which immigrant nationalities bear to the native-born whites of native parentage. Immigrants from the British Isles and English provinces of Canada, who have spoken the English language from birth, are held to be more like the native stock than are the immigrants from continental Europe. Of the latter, we regard the people of northwestern Europe as psychologically, culturally, and historically more nearly related to us than are the peoples of southern or of eastern Europe.

But little interpolation and re-arrangement are needed to convert this rough marking-scale of nationalities and races into a scientific scale of ten positions, as follows: 0, Native born of native white parents; 1, Native-born of foreign-born white parents; 2, Foreign-born, English-speaking; 3, Northwestern Europeans; 4, Southern Europeans and Latin-American whites; 5, Eastern Europeans; 6, All other whites; 7, Civilized yellow; 8, Civilized dark; 9, Uncivilized.

The principle on which this marking-scale is constructed for

the population of the United States may be applied to the population of any country by observing a general rule: Assign to the third position (numbered 2), all foreign-born elements speaking the language of the country to which they have come. Assign to the fourth position (numbered 3) the foreign-born that are psychologically, culturally, and historically more nearly related to the nationalities of position three than any remaining element is. Assign to position five (numbered 4) the foreign-born that are psychologically, culturally, and historically more nearly related to the nationalities of position four than any remaining element is, and so proceed.

In deciding what nationalities are "more nearly related" in the ways specified, we must not substitute the judgment of an individual observer for the consensus of popular opinion. In matters of this kind the "probable error" of thousands of individual judgments or markings making up the common opinion is extremely small.

There is, however, a scientific way of checking or testing the popular marking. It consists in breaking up each mark into a number of units, each of which stands for some specific and undisputed objective fact.

We may apply this test to the foregoing scale of ethnic positions, as follows:

Build up a composite statistical mark for "Americanism" by assigning one unit to each of the following facts, namely: (1) Parents native-born; (2) Self native-born; (3) Native language, that of the United States (English); (4) Reared under Celto-Teutonic traditions and cultural influences; (5) Reared under constitutional government; (6) Of European stock and reared under European civilization; (7) Belongs to race that has created an independent political state with a history; (8) Belongs to race that has created an ethical religion; (9) Belongs to race that has created a literature; (10) Belongs to race that has independently risen above barbarism; (11) Color lighter than yellow; (12) Color lighter than red; (13) Color lighter than brown; (14) Color lighter than black.

Assigning these units of marking to the respective elements

MARKING-SCHEME FOR NATIONALITIES

[illegible]

composing the population of the United States, we get the subjoined "Marking Scheme."

Passing from phenomena of population to phenomena of the social mind, we observe that simple marking-scales are daily used in connection with religion, conduct, ability, and education. The empirical gradings applied to these phenomena appear on examination to be not less natural or matter-of-course than the popular gradings of nationalities in respect of their resemblances.

If a nation has an established religion, it is assigned to the initial position. All other religions are looked upon as variate, kindred, or alien cults. In the United States we have no established religion, and we therefore assign to the initial position the denominations that in the earlier days of our history were proclaimed by a majority of the people to be orthodox, notwithstanding the fact that they were Protestant bodies, holding radical ideas on questions of church government. These denominations, after a time, divided on questions of doctrine, some of them becoming radical in belief, as well as in organization. Protestant bodies conservative in both belief and organization were formed at an early period, but for a time they were commonly regarded with disfavor. It was much later that the Roman Catholics gained a foothold outside of Maryland, and only recently have come the eastern Catholics and the non-Christian persuasions.

The essential relationships of these varying cults present a scheme or order closely correlated with the historical succession. The conservative Protestants, for example, are in matters of faith and worship more nearly related to the radical Protestants than the Roman Catholics are. The Roman Catholics, in turn, are more nearly related to the conservative Protestants than the Eastern Catholics or the non-Christian bodies are.

Accepting then the popular judgment upon these points of relationship, our marking-scale is as follows: Initial position or 0, Protestant radical A, including all religious bodies that are radical in both doctrine and organization; 1, Protestant radical B, including all Protestant bodies that are radical in organization only; 2, Protestant conservative; 3, Catholic A, including all

Roman Catholics; 4, Catholic B, including all Eastern Catholics; 5, Nominally Christian, but in many particulars reverting to the patriarchal Jewish type, namely, the Latter-Day Saints, or Mormons; 6, Non-Christian monotheistic, principally Jews; 7, All others.

This scale, expressive of popular judgment, may be tested, as we tested the scale for nationalities, by assigning unit marks to objective facts, one at least of which can be affirmed of all the religious groups, and all of which can be affirmed of at least one group. It will be found that the following facts meet these conditions, and we therefore assign a unit mark to each one: (1) Radical in doctrine; (2) Radical in government; (3) Protestant; (4) Holding to Western European Christian tradition; (5) Holding to general European Christian tradition; (6) Holding to generic Christian tradition; (7) Monotheistic; (8) Theistic.

Assigning these units of marking to the various groups of religious bodies found in the United States, we get the following ratings:

MARKING-SCHEME FOR RELIGIONS

	Radical in Doctrine	Radical in Government	Protestant	Holding to Western European Christian Tradition	Holding to General European Christian Tradition	Holding to Generic Christian Tradition	Monotheistic	Theistic	Total Mark	Position
Protestant radical A.	1	1	1	1	1	1	1	1	8	0
Protestant radical B.	..	1	1	1	1	1	1	1	7	1
Protestant conservative	1	1	1	1	1	1	6	2
Catholic A	1	1	1	1	1	5	3
Catholic B	1	1	1	1	4	4
Mormon	1	1	1	3	5
Jewish	1	1	2	6
All others	1	1	7

The marking-scale applied to conduct is an inheritance from the earliest days of civilization, and is substantially the same throughout the civilized world. Its gradations designate objectionable, rather than commendable, qualities of behavior. To

the initial position 0, we assign the inoffensive and law-abiding; to 1, we assign the minor misdemeanants; to 2, the major misdemeanants; to 3, the felons whose offenses are not capital; to 4, the felons whose offenses are capital.

On the marking-scale for ability, we assign to the initial position those who are competent to work under direction; to 1, we assign those who are capable of self-direction in a small way; to 2, those who can supervise in accordance with general instructions provided for them; to 3, those competent to exercise responsible direction, a group which includes the great majority of land-owning farmers, business men, and professional men; and to 4, we assign the directors of directors, namely, the greater organizers, statesmen and financiers, the greater discoverers and inventors, and the greater professional men and philosophers.

The marking-scale for education, in like manner, is simple, and is the same for all occidental countries.

The groups or classes assigned to its successive positions are: 0, the illiterate; 1, those educated not beyond the grammar-school grades; 2, those educated not beyond the high-school grades; 3, those educated not beyond the college grades; 4, those educated in university grades.

Assignments to positions on the conduct-scale and education-scale are matters of record, and usually it is not necessary to go behind the returns. Each grade, however, and each grade of ability admits of indefinite subdivision, and assignments to the intermediate positions must be made by experts. The probable error of their gradings is the measure of the scientific value of the marks. The most admirable example of the possibilities of such gradings that we have is found in Professor Cattell's *Directory of American Men of Science*.

We may pass now to the inquiry: What sociological problems may be stated and solved in statistical terms with the aid of these social marking-scales?

The answer turns upon a technical point, which may be stated in the form of a preliminary question, namely, Are the "distances" or "steps" between successive positions on any given

marking-scale always equal, or may they be regarded as equal for the purposes of sociological induction? May we assume, for example, that the step from "native-born of native parents" to "native-born of foreign parents" is neither longer nor shorter than the step from "native-born of foreign parents" to "foreign-born, English-speaking"? Or may we assume that the step from "Protestant radical B" to "Protestant conservative" is equal to the step from "Protestant conservative" to "Catholic A"? If these steps or any others on the same scale are unequal, we may still find the scale worth while, and from the data to which it is applied we may draw conclusions within those limits that have been indicated by Galton, Bowley, and Thorndike. If, however, the steps are equal, or may legitimately be treated as if they were equal, the use of the scale is greatly extended.

What then are the facts? Is it possible for any individual or group of experts to prove, or for the general public to "feel sure," that the "steps" in these various marking-scales are distinguishable in magnitude? If it is alleged that they are distinguishable, and if the differences can be specified, we can overcome the difficulty, at least in part, by weighting our marks. If, on the other hand, it is admitted that no one can say whether the "steps" are equal, or not, a fact of great consequence to scientific sociology is thereby established.

The significance of differences and inequalities among factors of a social population lies in the fact that they are reflected in social feeling and impair the sense of social solidarity, thereby affecting co-operative efficiency. Race conflicts, religious wars, and persecutions, and class conflicts, are expressions of a conscious antagonism which depends for its existence upon a consciousness of difference. If there is no feeling of difference, there is no prejudice; and if differences are too slight or too vague to be detected, the feeling of difference cannot arise.

Therefore, differences among social factors which are not perceived or felt by the social units themselves are without effect upon the important phenomena of prejudice, the consciousness of kind in general, and the sense of social solidarity. For the purposes of sociological theory they may be neglected.

Similar reasoning applies to inequalities between any two differences and any other two. The difference between one nationality and another may be obvious, but it may be impossible to perceive or even to feel that the difference between nationalities C and D is greater than the difference between nationalities A and B. This is precisely the kind of difference that we contemplate when we ask whether the "steps" in any given social marking-scale are equal.

We conclude, then, that when a social marking-scale has been so constructed that it is not possible for the expert or the public to say whether the successive "steps" of the scale are equal, or not, any absolute difference which may undetected exist is negligible for the purposes of sociological theory. The "steps" may be assumed to be equal.

Adopting this conclusion, we proceed to show how the social marking-scales may be used in the statistical analysis of sociological problems.

The following table presents the census data on population for the years 1890 and 1900, grouped according to the marking-scale for nationalities and races (Table I).¹

Plotting the data of this table as a surface of magnitude, according to Galton's method,² we get the rough curves of diminishing magnitude, or sub-homogeneity, shown in Figs. 1 and 2. Plotting as a surface of frequency, we get the rough curves of frequency shown in Figs. 3 and 4.

These curves show in a rather striking way the more significant relations of our numerous ethnic elements to one another and to the total population, and in particular the extent of departure from an ideal homogeneity. The true measures of this

¹ I am indebted to Mr. F. Stuart Chapin for the computations involved in preparing the tables for this paper.

² *Natural Inheritance*, pp. 37, 38. Galton uses the terms "polygon of distribution" and "curve of distribution" to distinguish from "polygon of frequency" and "curve of frequency," the forms obtained by plotting magnitudes on the vertical y , instead of on the horizontal x . Inasmuch as the "curve of frequency" is also a "curve of distribution," it is more exact to use the term "surface or curve of magnitude" when we plot magnitudes on the vertical and frequencies on the horizontal, and the term "surface or curve of frequency" when we plot magnitudes on the horizontal and frequencies on the vertical.

TABLE I
RESEMBLANCE POSITIONS OF NATIONALITIES AND RACES: CONTINENTAL
UNITED STATES

NATIONALITIES AND RACES	POSITIONS	1890		1900	
		Number	Percentage	Number	Percentage
Native-born of native white parents	0	34,468,602	.547	41,356,097	.541
Native-born of foreign-born white parents	1	11,503,675	.182	15,687,322	.205
Foreign-born, English speaking	2	3,807,337	.060	3,581,113	.047
English		908,141		842,078	
English Canadian		678,442		785,958	
Scotch		242,231		233,977	
Welsh		100,079		93,682	
Irish		1,871,509		1,618,567	
Australian		5,984		6,851	
Not Specified		951		
Foreign-born, North-western Europe	3	4,468,502	.071	4,911,163	.064
French		113,174		104,341	
French Canadians		302,496		395,297	
Belgians		22,639		29,804	
Germans		2,784,894		2,668,990	
From Luxembourg		2,882		3,041	
German Poles		150,232	
Hollanders		81,828		105,049	
Danes		132,543		154,284	
Swedes		478,041		573,040	
Norwegians		322,665		336,985	
Swiss		104,069		115,851	
Austrians		123,271		276,249	
Southern Europeans and Latin Americans	4	313,955	.005	668,119	.009
Spaniards		6,185		7,072	
Portuguese		15,996		30,618	
Italians		182,580		484,207	
Greeks		1,887		8,564	
Mexicans		77,853		103,410	
Central Americans		1,192		3,901	
South Americans		5,006		4,761	
Cubans		11,153	
West Indians		23,256		14,433	
Eastern Europeans	5	510,623	.008	1,038,019	.014
Russians		182,644		424,096	
Finns		62,811	
Russian Poles		154,424	
Austrian Poles		58,503	
Poles		147,440		
Unknown Poles		20,351	

TABLE I—*Continued*

NATIONALITIES AND RACES	POSITIONS	1890		1900	
		Number	Percentage	Number	Percentage
Hungarians.....		62,433		145,802	
Bohemians.....		118,106		156,991	
Roumanians.....			15,041	
All other whites*.....	6	32,376	.0005	35,319	.0004
Turks.....		1,839		9,933	
Europeans not specified..		12,579		2,263	
Born at sea.....		5,533		8,220	
Other countries.....		479		2,558	
Atlantic Islands.....		9,539		9,784	
Africa.....		2,207		2,552	
Civilized Yellow.....	7	111,240	.002	118,812	.002
Chinese.....		106,688		81,827	
Japanese.....		2,202		25,077	
Asian (not specified).....		2,260		11,908	
Civilized Dark.....	8	7,553,794	.120	8,869,342	.116
Civilized Negroes.....		7,488,676		8,749,735	
Civilized Indians.....		58,806		107,706	
Hindoos.....		2,143		2,050	
Sandwich Islanders.....		1,304		
Pacific Islanders.....		2,065		2,049	
Uncivilized and otherwise not specified.....	9	180,447	.003	129,518	.002
Uncivilized Indians.....		189,447		129,518	
Totals.....		62,959,551	.100	76,394,824	.100
Sub-homogeneity.....		1.579		1.552	
Sub-homogeneity from percentages.....		.01579		.01552	

* The group "All Other Whites" as here made up should in strict accuracy be distributed. It should include Turks, Persians, Armenians, African whites, not descended from European stock, etc., while "Europeans not Specified," "Born at Sea," and from "Atlantic Islands," should be assigned to preceding groups. The labor of ascertaining what those assignments should be was not worth while, the entire group constituting only .0004 per cent. of our total population.

departure are the positions of the medians and the quartiles. Notwithstanding the magnitude of the immigration which we have received within the last twenty-five years, nearly three-fourths of the entire population of continental United States is native-born, and more than one-half is native-born of native parents. Nearly 80 per cent. has been English-speaking from birth. These facts of course are disclosed by the census data independently of the marking-scale here used. But the scale and the plotted curves reveal one tendency that, without their aid,

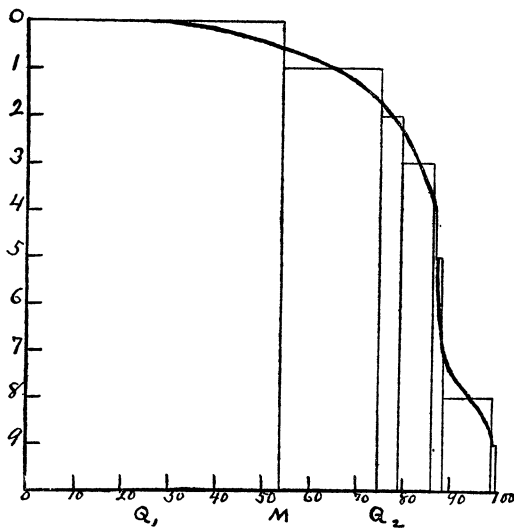


FIG. 1.—Nationalities: Magnitude Curve of Resemblance Positions, 1890

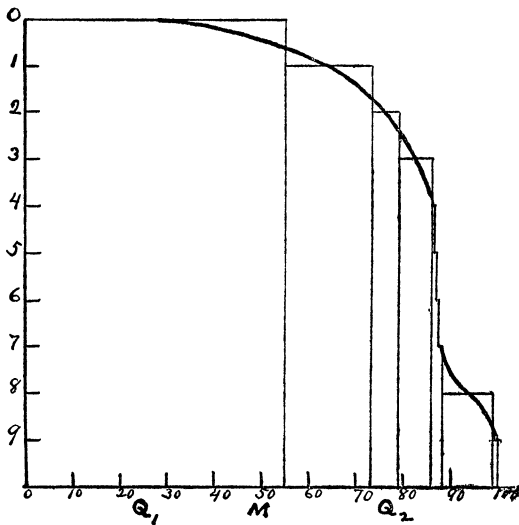


FIG. 2.—Nationalities: Magnitude Curve of Resemblance Positions, 1900.

would not easily be discovered, namely, the curve of 1900 is slightly less irregular than the curve of 1890. Very slightly, but perceptibly, it is nearer the straight line which would express the divergence of nationalities from the native-born of native par-

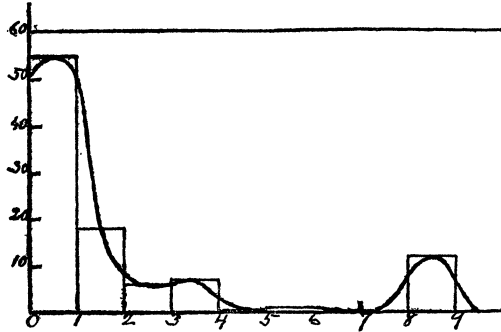


FIG. 3.—Nationalities: Frequency Curve of Resemblance Positions, 1890.

ents, if each successive step of divergence were equal to the step preceding, and if each nationality were represented in our population by the same number of individuals. This circumstance suggests a question: Is our population becoming slightly less

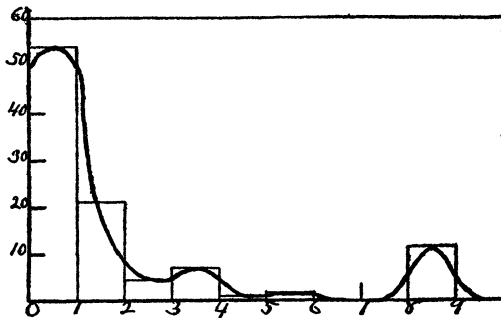


FIG. 4.—Nationalities: Frequency Curve of Resemblance Positions, 1900.

heterogeneous, contrary to the prevailing belief? We shall presently see that it is.

The curves of frequency show the positions of the modes. A multi-modal curve, as all statisticians understand, indicates a compound group, or one made up of two or more kinds of units.

The American population is racially compound, and the actual gulf between the native-born whites and the negro element is clearly shown by the primary mode at 0 (Figs. 3 and 4), and

TABLE II

RESEMBLANCE POSITIONS OF RELIGIOUS BODIES: CONTINENTAL UNITED STATES

RELIGIOUS BODIES	POSITION	1890		1906	
Protestant radical A.....	0		1.3		1.3
Adventist.....		0.3		0.3	
Christians.....		0.5		0.6	
Christian Science.....		0.0		0.3	
Unitarians.....		0.3		0.2	
Universalist.....		0.2		0.2	
Protestant radical B.....	1		51.8		45.7
Baptist.....		18.0		17.0	
Congregationalist.....		2.5		2.1	
Disciples.....		3.1		3.5	
Dunkers.....		0.4		0.3	
Evangelical bodies.....		0.6		0.5	
Friends.....		0.5		0.3	
German Evangelical.....		0.9		0.9	
Independent.....		0.1		0.2	
Mennonites.....		0.2		0.2	
Methodist.....		22.3		17.5	
Reformed.....		1.5		1.4	
United Brethren.....		1.1		0.9	
Other Protestant.....		0.6		0.7	
Protestant Conservative.....	2		14.8		14.7
Lutheran.....		6.0		6.4	
Presbyterian.....		6.2		5.6	
Protestant Episcopalian.....		2.6		2.7	
Catholic A.....	3		30.3		36.7
Roman Catholic.....		30.3		36.7	
Catholic B.....	4		0.0		0.4
Eastern Orthodox.....		0.0		0.4	
Mormon Christian.....	5		0.8		0.8
Latter-Day Saints.....		0.8		0.8	
Non-Christian A.....	6		0.6		0.3
Jewish.....		0.6		0.3	
Non-Christian B.....	7		0.3		0.2
All non-Christian religious bodies not included under non-Christian A.....		0.3		0.2	
Sub-homogeneity from percentages.....		1.833		1.953	

the secondary mode at 8. Moreover, a tendency toward the formation of lesser modes among the foreign-born, slightly manifest in 1890, is distinctly more marked in 1900.

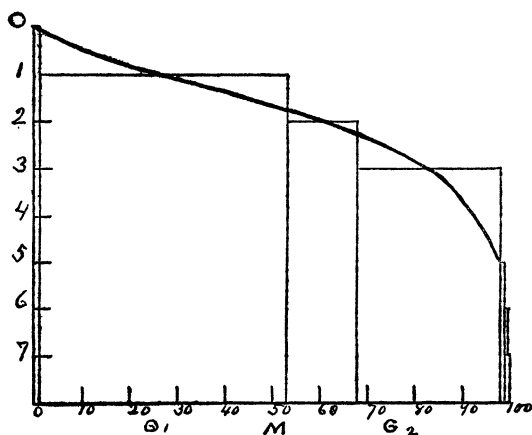


FIG. 5.—Religions: Magnitude Curve of Resemblance Positions, 1890.

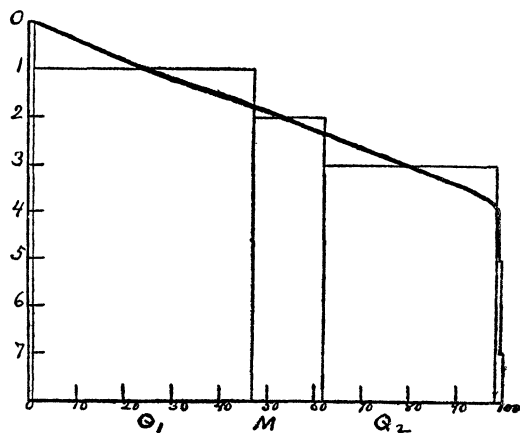


FIG. 6.—Religions: Magnitude Curve of Resemblance Positions, 1906.

Table II presents the census data on religious bodies for the years 1890 and 1906, grouped according to the marking-scale for religion.

Plotting the data of this table as surfaces of magnitude, we get the rough curves of sub-homogeneity shown in Figs. 5 and

6. Plotting as surfaces of frequency we get the rough curves of frequency shown in Figs. 7 and 8.

If marking-scales of social phenomena, and the curves of magnitude and of frequency which we are able to plot by means

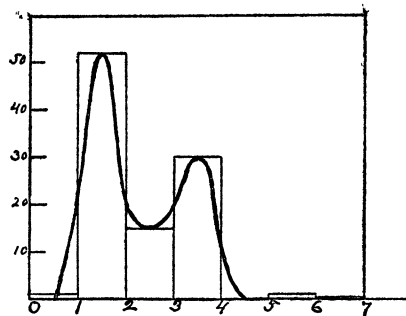


FIG. 7.—Religions: Frequency Curve of Resemblance Positions, 1890.

of them, had no other use than that of enabling us to form approximately correct notions about the composition of a population, and the make-up of its social mind, they would be serviceable enough to repay the labor involved in applying them. They

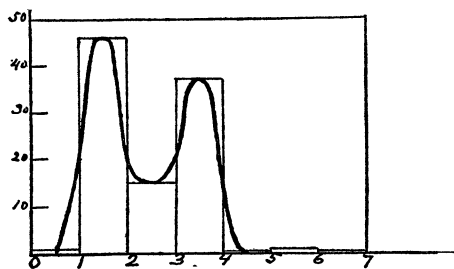


FIG. 8.—Religions: Frequency Curve of Resemblance Positions, 1906.

have, however, a further and more important use, which must now be indicated.

The question has been raised whether the American population is becoming more or less heterogeneous in ethnic composition. Not less interesting are the questions: Are we becoming more heterogeneous or less heterogeneous in religious belief?

Are we becoming more unlike or less unlike in conduct? Are we becoming more unequal or less unequal in educational attainments?

But to ask these questions is to raise yet another which is preliminary to them all, namely: Is it in any case possible to measure relative heterogeneity? Given two compounds, is it possible to determine how much more compounded one is than the other?

I do not know what attempts may have been made to answer this question in its generalized form, but I believe that no method has hitherto been suggested, or at any rate used, for measuring the sub-homogeneity of mixed populations, or of the social mind. I venture therefore to suggest that the problem admits of solution and to offer a formula for finding a numerical measure or coefficient of heterogeneity, or, as I think we may more accurately say, of sub-homogeneity.

When we plot the surface of magnitude (the polygon of distribution as defined by Galton), we measure magnitudes, or we determine successive positions, by vertical distances from a baseline. Since equal magnitudes must be measured, and positions that indicate equal degrees of relationship must be determined, by equal vertical distances, the plotted points indicating such equal measures must lie in a horizontal straight line. Therefore, the horizontal straight line is the graphic expression of equality or of homogeneity. Departures from equality, or degrees of sub-homogeneity, are best plotted as minus or negative magnitudes, graphically expressed by points placed at proper distances below a horizontal line, which itself expresses the ideal homogeneity from which departures are supposed to be made.

Let us then suppose that we have to describe a group of human beings, twelve in number, of whom three are unobjectionable in conduct; three are vicious persons, or minor misdemeanants; two are petty criminals or major misdemeanants; two are felons, not capital; and two are felons, capital. The group as a whole is sub-homogeneous. Plotting the data, we get the arrangement shown in Fig 9.

Imagine now, that by expending one unit of some kind of

effort, we could lift any one vicious person up to the level of the men whose conduct is unobjectionable; that by expending two units of the same kind of effort we could lift any one major misdemeanor to the same standard level; that by expending three units of the same kind of effort, we could lift any one of the minor felons, and by expending four units of the same kind of effort we could lift any one of the major felons, to our standard level. Then, by expending $1 \text{ unit} \times 3 + 2 \text{ units} \times 2 + 3 \text{ units} \times 2 + 4 \text{ units} \times 2$, or 21 units in all, or 1.75 units per capita, for the whole group of twelve persons, we should convert the entire sub-homogeneous group of twelve persons into a group

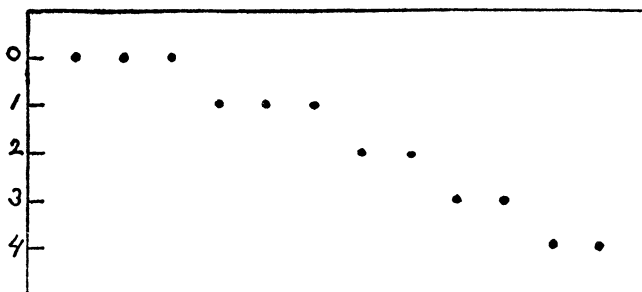


FIG. 9.—Measurement of Sub-Homogeneity.

perfectly homogeneous in respect of a standardized conduct.

Students of the physical sciences who are accustomed to measure physical phenomena of every description by the number of units of effort, or of energy, necessary to transform them from one state into another, will assent to the proposition that if we could thus actually transform any heterogeneous group of human beings into a homogeneous group, the number of units of effort necessarily expended in the process could be taken as an accurate measure of the total sub-homogeneity of the original group, and that the total number of units of effort so expended, divided by the total number of individuals in the transformed group, could be taken as a measure of the per-capita degree of sub-homogeneity of the original group.

Is it legitimate to conceive of any heterogeneous group as ideally transformable by such a procedure, and then to assume

that we may measure its sub-homogeneity by (1) multiplying each successive numerical mark on a marking-scale of resemblance positions by the "frequency" or number of individuals assigned to that position, (2) obtaining the sum of the products, and (3) dividing it by the whole number of individuals in the group or population?

It is, I think, an adequate and satisfactory answer to this question to observe that the conception and the assumption are legitimate, if mankind is warranted in believing that by an expenditure of educational and reformatory effort, it can standardize knowledge and conduct, and can assimilate alien habits and ideals to prevailing or national types. If the validity of this pragmatic belief be conceded, there can be no objection to conceiving of an average and abstract unit of standardizing effort, practically unchanging throughout the same group or population, living under practically constant conditions.

If so much be granted, we may write the formula for measuring sub-homogeneity as follows:

Designate positions on the marking-scale by the numerals, 0, 1, 2, 3, 4, n .

Designate frequencies by $K_0, K_1, K_2, K_3, K_4, \dots K_n$

Designate total individuals, or population by P .

Designate per-capita degree of sub-homogeneity by S .

Then:

$$S = \frac{K_1 + 2K_2 + 3K_3 + \dots + NK_n}{P}.$$

Applying this formula to the data presented in the tables of ethnic distribution by resemblance positions, we obtain for 1890 the coefficient .01579 as our expression for the per-capita degree of sub-homogeneity, and for 1900 the coefficient .01552. Applying the formula to the data presented in the tables of religious distribution by resemblance positions, we obtain for 1890 the coefficient 1.833, and for 1906 the coefficient 1.953. These coefficients indicate that in ethnic composition the people of the United States are now becoming, contrary to the prevailing impression, slightly more homogeneous, but that in religious persuasion they are still becoming slightly more heterogeneous.